# 📌 Explanation of Each Metric Example for Logistic Regression

This document provides real-world examples demonstrating when to use specific metrics in Logistic Regression models.

## 1️⃣ Accuracy - Spam Email Classification ✉️

\*\*When to Use Accuracy?\*\*  
✅ If the dataset is balanced (i.e., roughly equal numbers of each class).  
✅ If all misclassifications have similar consequences.

\*\*Example Scenario:\*\*  
An email service wants to classify emails as spam (1) or not spam (0).  
- \*\*Independent Variables (X):\*\* Email content, sender, number of links, etc.  
- \*\*Target Variable (Y):\*\* 1 (Spam), 0 (Not Spam).  
- \*\*Model:\*\* Logistic Regression.

📉 \*\*Result:\*\*  
\*\*Accuracy = 93%\*\* → The model correctly classifies \*\*93% of emails\*\* as spam or not spam.

🔹 \*\*Why Accuracy?\*\*  
Since spam detection typically has \*\*balanced classes\*\*, accuracy is a reliable metric.

## 2️⃣ Precision - Fraud Detection in Transactions 💳

\*\*When to Use Precision?\*\*  
✅ If false positives are more dangerous than false negatives.  
✅ Best for fraud detection, medical tests, or financial risk models.

\*\*Example Scenario:\*\*  
A bank wants to detect fraudulent transactions (1) versus genuine transactions (0).  
- \*\*Independent Variables (X):\*\* Transaction amount, location, device used.  
- \*\*Target Variable (Y):\*\* 1 (Fraud), 0 (Legit).  
- \*\*Model:\*\* Logistic Regression.

📉 \*\*Result:\*\*  
\*\*Precision = 87%\*\* → Of all transactions flagged as fraud, \*\*87% were actually fraudulent\*\*.

🔹 \*\*Why Precision?\*\*  
Since \*\*false positives (flagging a legit transaction as fraud) cause inconvenience\*\*, we prefer high precision.

## 3️⃣ Recall - Disease Diagnosis in Healthcare 🏥

\*\*When to Use Recall?\*\*  
✅ If false negatives are worse than false positives.  
✅ Used in healthcare, security systems, and risk management.

\*\*Example Scenario:\*\*  
A hospital wants to classify whether a patient has a certain disease (1) or not (0).  
- \*\*Independent Variables (X):\*\* Age, symptoms, test results.  
- \*\*Target Variable (Y):\*\* 1 (Has disease), 0 (Healthy).  
- \*\*Model:\*\* Logistic Regression.

📉 \*\*Result:\*\*  
\*\*Recall = 92%\*\* → Of all patients who actually have the disease, \*\*92% were correctly identified\*\*.

🔹 \*\*Why Recall?\*\*  
Since \*\*missing a disease case (false negative) is more dangerous than a false alarm\*\*, we prioritize high recall.

## 4️⃣ ROC-AUC Score - Stock Price Movement Prediction 📈

\*\*When to Use ROC-AUC?\*\*  
✅ If you want to measure the model's ability to distinguish between classes.  
✅ Best for evaluating performance on imbalanced datasets.

\*\*Example Scenario:\*\*  
A trader wants to predict whether a stock will go up (1) or down (0) the next day.  
- \*\*Independent Variables (X):\*\* Opening price, trading volume, technical indicators.  
- \*\*Target Variable (Y):\*\* 1 (Stock goes up), 0 (Stock goes down).  
- \*\*Model:\*\* Logistic Regression.

📉 \*\*Result:\*\*  
\*\*ROC-AUC = 0.89\*\* → The model has an \*\*89% chance\*\* of correctly distinguishing between rising and falling stocks.

🔹 \*\*Why ROC-AUC?\*\*  
Since stock movements can be \*\*imbalanced (e.g., stocks go up more often than down)\*\*, ROC-AUC is a great measure of performance.

## 📌 Final Takeaways

✅ \*\*Use Accuracy for spam detection\*\* → When classes are balanced.  
✅ \*\*Use Precision for fraud detection\*\* → When false positives are costly.  
✅ \*\*Use Recall for disease diagnosis\*\* → When missing a case is dangerous.  
✅ \*\*Use ROC-AUC for stock prediction\*\* → When dealing with imbalanced datasets.